

Performance evaluation of feature selections on some ML approaches for diagnosing the narcissistic personality disorder

Heni Sulistiani^{1,2}, Admi Syarif³, Kurnia Muludi³, Warsito³

¹Doctoral Program in Faculty Mathematics and Natural Science, Lampung University, Lampung, Indonesia

²Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Lampung, Indonesia

³Department of Computer Science, Faculty of Mathematics and Natural Science, Lampung University, Lampung, Indonesia

Article Info

Article history:

Received May 12, 2023

Revised Oct 11, 2023

Accepted Oct 19, 2023

Keywords:

Feature selection

Machine learning

Mental health

Narcissistic

Personality disorder

ABSTRACT

Narcissistic personality disorder (NPD) is a personality disorder that affects various aspects of life, including relationships, employment, school, and finances. Persons with NPD usually feel unhappy and disappointed when no one helps them and is not praised for their achievements. Diagnosing narcissism is generally done using a screening test that consumes time and costs a lot. This research aims to evaluate the performance of several feature selection (FS) approaches on machine learning (ML) techniques (support vector machine (SVM), random forest classifier (RFC), and Naive Bayes). Three scenarios of FS (all features, the information gain technique and the gain ratio (GR) feature technique) are used for each ML method. Several experiments using the benchmark narcissistic disorder dataset have been done. It adopts the *k*-fold cross-validation (10-fold cross-validation) strategy. We evaluate the method's performance by measuring its accuracy, error rate, and processing time. It is shown that the RFC GR strategy gives the best performance with an accuracy of 100%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Admi Syarif

Department of Computer Science, Faculty of Mathematics and Natural Science, Lampung University

Jl. Sumantri Brojonegoro No. 1, Bandar Lampung, Lampung, Indonesia

Email: admi.syarif@fmipa.unila.ac.id

1. INTRODUCTION

The problem of mental health is a serious issue in modern society. It usually refers to a person's mood, thoughts, and behavior [1]. One of the leading causes of suicide is poor mental health [2]. A personality disorder is one type of mental health disorder that interferes with ways of thinking, understanding situations, and relating to others. A person with a personality disorder can self-destruct in any situation [3]. There are several types of personality disorders, such as narcissistic, paranoid, schizoid, antisocial, borderline, and so on. A 2019 World Health Organization (WHO) survey shows nearly a billion people have a mental disorder [4]. People with mental disorders are frequently stigmatized in the community, which can have a negative impact on them [5].

Narcissism is a common mental disorder [6]. The characteristics of narcissistic personality disorder (NPD) narcissism are grandiosity, a need for praise, and a lack of empathy [7]. It is becoming more common in everyday life, particularly in social media. Those suffering from NPD frequently exhibit an inability to maintain relationships and their jobs. The NPD causes problems in many areas of life, including relationships, work, school, and finances. People with NPD may become dissatisfied and disillusioned when they do not receive special favors or adoration, they believe they are entitled to. They may not enjoy being around them because their relationships are not rewarding [8], which may cause them to act violently and aggressively [9].

Early personality disorder diagnosis is critical for understanding people's psychological conditions [10]. As a result, we can take steps to help people with personality disorders [11]. It is done based on appearance or behavior, which is usually a complex and difficult task. It requires the medical science of psychology [12]. It usually employs test screening. However, it would be costly and time-consuming for a large population. Furthermore, diagnostic procedures have the unintended consequence of discouraging healthy people from participating. Psychological problems consequently frequently go unrecognized or addressed. Machine learning (ML) is one method for predicting mental health disorders quickly and accurately. In recent years, ML techniques have been used in a variety of medical research projects, particularly in biomedical and neuroscience, to gain a better understanding of mental health issues [13].

ML enables computers and other computing systems to autonomously learn from the past and advance without explicit human programming [11]. ML is based on the creation of computer programs that can access data and figure things out for themselves. It is very effective in the healthcare industry, where there is a lot of data. In that case, the prediction model produced will be superior, free of human error, and will reduce the time required for diagnostics. ML is a significant advancement in computer science and data processing techniques that can be used to improve almost any service [14]. This allows for the formation of a pattern in the data, allowing for easier and more accurate identification and prediction. It is considered to be a highly useful tool for predicting mental health [15]. ML requires input in the form of features or variables. One of the most important issues is determining which feature will lead us to the best solution. Datasets used in the ML process typically contain redundant and irrelevant features, cannot improve accuracy [16], have no effect on the learning model [17], and may even degrade the learning model's performance [16]. As a result, relevant features must be chosen.

Numerous studies on feature selection (FS) techniques have been conducted recently to create a dataset with the most pertinent features for the best model performance [18] and reduce computation time [19]. FS is typically used to select only efficient features based on the given input by reducing noisy data, which aids in the identification of the application [20]. There have been many reports in previous research on FS methods that affect improving ML performance, including relief [21], minimum-redundancy-maximum-relevance FS (m-RMR) [22], information gain [23]-[25], and gain ratio (GR) [23], [26], [27]. To our knowledge, no research has reported the best method for each ML approach. Each method employs a distinct technique or strategy for identifying the best and most relevant features. There is no research that compares ML performance or discusses the use of FS in NPD cases.

In this study, we would like to compare the performance of various FS approaches (all features, information gain, and GR) against some ML methods (support vector machine (SVM), random forest classifier (RFC), and Naive Bayes). In this study, features are chosen when the threshold value is greater than 0.05. It is frequently required and used in studies that require good accuracy at a lower level [28], [29]. The research findings will provide an overview of the ML approach's ability to predict NPD using features generated by FS techniques.

The rest of this paper is organized as follows: section 2 provides a brief overview of the research method. In section 3, we present the experiment results as well as the comparison classification and FS methods. Finally, section 4 describes the conclusion, which demonstrates the remarkable effectiveness of the approaches.

2. METHOD

This section outlines the procedures followed when doing research. The research stage begins with determining the problem, followed by data collection. Before the data is entered into the model, it is necessary to do pre-processing of the data. At this stage, the data is cleaned and adjusted to be processed for the next step (feature selections using information gain and GR). The next stage is that the selected features will be processed into the ML method (SVM, RFC, and Naive Bayes). It consists of several steps, as illustrated in Figure 1.

2.1. Dataset

We have selected 8376 data from the data collecting process, with an average age of 14 to 50, 44 features, and 1 class label. The data used in this research are obtained from Open Psychometrics (<https://www.kaggle.com/datasets>). The dataset consists of 5330 data with the class label of "yes" and 3046 with the class label of "no." Utilizing feature selection and preprocessing approaches, the data have been cleaned and reduced to the most important attributes. Three steps make up the method: feature selection, preprocessing, and applying ML to the prediction process. The details of 44 features of this study are listed in Table 1.

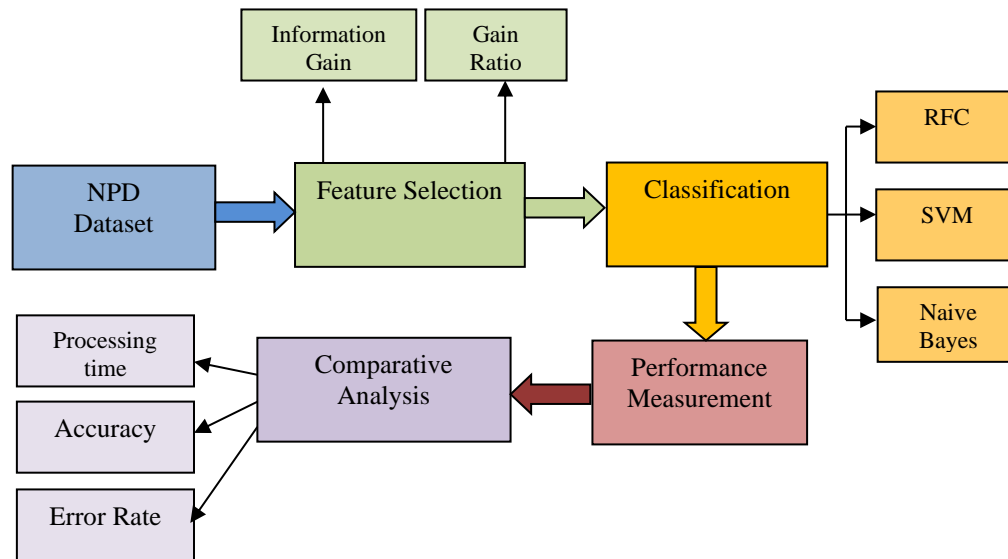


Figure 1. The design of the proposed method

Table 1. Features of dataset

Feature	Description
f01	Gender
f02	Age
f03-f42	Statements 1–40
f43	Elapse
f44	Score
f45	Class

2.2. Feature selection

The source dataset often consists of various features, some of which may or may not be important for categorization [30], [31]. Unimportant features that depend on other attributes reduce prediction accuracy. A feature selection strategy must be used to overcome this and decrease the feature's dimension. There are several methods in feature selection, namely filter, wrapper, and hybrid methods. The method most often used for feature selection is the filter method [32]. The Information gain and GR methods are used in this work, which helps to find the feature subsets. The techniques are one of the popular filter models [29]. Each feature in the dataset was counted, selected, and defined using a value limit known as the threshold (cutoff). This research uses a threshold >0.05 .

2.2.1. Information gain

Information gain is the change in class entropy from a previous state to a state when an attribute value occurs. It is applied here to demonstrate how features are pertinent. Decision tree induction is the foundation of this method. Information gain is used as a criterion for choosing attributes. The information gain method has a faster time in the feature selection process than other methods [33]. The features with the most information will be ranked highly in this method; otherwise and low. It is computed using the following steps:

Step 1: step one involves computing entropy (H) before observing attribute A and class C .

Step 2: once attribute A has been observed, calculate the entropy. After seeing features that are a part of subsets of the main data set, this calculates the entropy. The first two phases are extremely important since they supply the entropies needed in the next phase to acquire information gain.

Step 3: information gain calculation. The difference between the Entropy before observing characteristic A and information gain of attribute A .

The final feature set that will be utilized for classification is chosen after computing the gain from each feature and setting a threshold.

2.2.2. Gain ratio

After dividing the data, the entropy value of the probability distribution subset is calculated using the GR, which normalizes the information gain acquired [21], [34]. When selecting features, the GR

considers the dataset's number and size [35]. The GR modifies the information gain, which lessens its bias. The GR chooses an attribute based on the number and size of branches. By accounting for the inherent information of a split, it corrects information gain.

2.3. Classification

A classification is a form of data mining technique that is currently popular. It functions similarly to other methods like decision trees and neural networks. These strategies use various methods to assess the available data to produce their prediction.

After the feature selection stage and the features that affect the class label based on ranking are obtained, the next stage is classification. Furthermore, the accuracy, error rate, and time required in the classification process are compared using the method of SVM, RFC, and Naive Bayes. The classification model will be validated using k -fold cross-validation. The cross-validation method is commonly used for training sets [36]. Figure 2 shows the k -fold cross-validation.

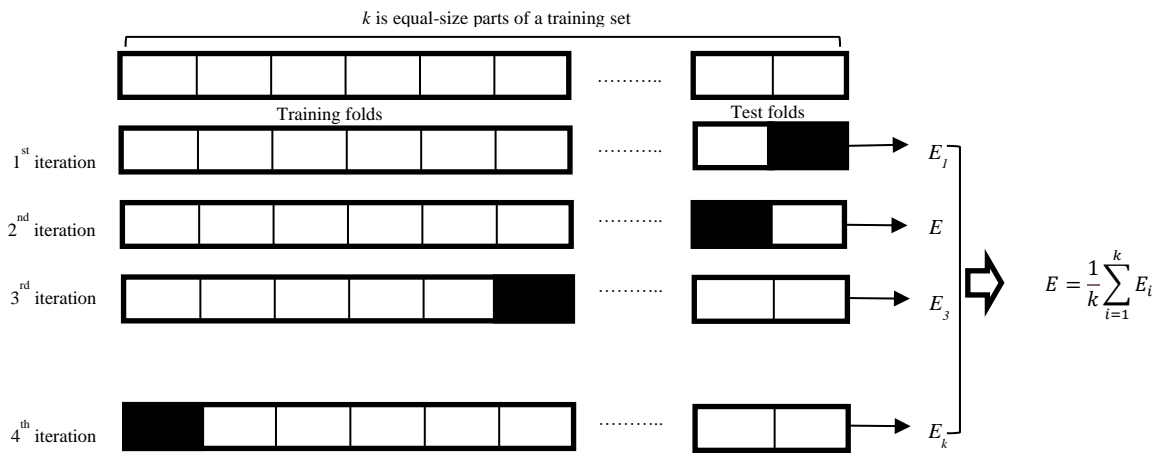


Figure 2. The procedure of k -fold validation

2.4. Performance measurement

This study examines how the confusion matrix might be used to gauge accuracy and mistake rate. A confusion matrix of size $n \times n$ coupled to a classifier, where n is the total number of classes, displays the anticipated and actual categorization. The confusion matrix for $n = 2$ is shown in Table 2.

Table 2. Confusion matrix		
	Predicted positives	Predicted negatives
Actual positives instances	Number of true positives instances (TP)	Number of false negatives instances (FN)
Actual negatives instances	Number of false positives (FP)	Number of true negatives instances (TN)

Calculating prediction accuracy and classification error is another method for evaluating and comparing classifiers. Both values can be obtained from the confusion matrix Table 2 and calculated using (1) and (2).

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

Meanwhile, to calculate the error, the following equation is used.

$$Missclassification (Error) Rate = \frac{(FP+FN)}{(TP+TN+FP+FN)} \quad (2)$$

3. RESULTS AND DISCUSSION

In the pre-processing stage of data mining, features are selected from the initial attributes through feature selection. In this research, we use information gain and GR feature selection techniques to determine the number of features based on weight values. The first step in the method for choosing features is calculating the weight of each feature. Next, features will be ranked based on the weight value.

In this research, we use three scenarios for performance evaluation of the classification algorithm including accuracy, error rate, and computational time. The first scenario: use all features. Second scenario: feature selection is carried out using information gain techniques. Third scenario: feature selection is carried out using the GR technique. The main aim was to perform a comparative analysis of the use of all features and two different feature selection techniques. In addition, to identify the best feature selection techniques that recommend the most relevant features. The results of feature selection can be seen in Table 3.

Table 3. Selected features

Information gain		GR	
Feature	Ratio	Feature	Ratio
f44	0.971	f44	1.000
f42	0.170	f42	0.173
f32	0.152	f32	0.160
f11	0.150	f29	0.157
f14	0.143	f38	0.155
f35	0.141	f37	0.155
f34	0.132	f34	0.151
f29	0.132	f11	0.150
f15	0.127	f06	0.144
f37	0.127	f14	0.140
f09	0.119	f35	0.138
f10	0.119	f15	0.136
f08	0.115	f09	0.134
f41	0.115	f27	0.125
f12	0.111	f08	0.120
f33	0.111	f10	0.119
f38	0.111	f22	0.118
f07	0.107	f41	0.116
f27	0.098	f25	0.115
f22	0.098	f12	0.111
f06	0.096	f33	0.110
f36	0.095	f07	0.109
f20	0.090	f40	0.105
f31	0.090	f04	0.102
f03	0.085	f36	0.101
f28	0.085	f20	0.099
f21	0.079	f30	0.092
f18	0.078	f21	0.091
f04	0.077	f31	0.089
f25	0.071	f03	0.087
f30	0.067	f28	0.085
f40	0.066	f17	0.820
f39	0.066	f18	0.077
f17	0.062	f23	0.074
f23	0.058	f05	0.072
f13	0.057	f39	0.072
f16	0.057	f16	0.068
		f13	0.065

From the Table 3 shows that the information gain technique produces 37 selected features, while the GR technique produces 38 features. From the information gain technique, there are seven features with a weight value of less than 0.05, including f26, f05, f19, f24, f02, f01, and f43. Meanwhile, from the GR technique, there are six features that with a weight value of less than 0.05, including f26, f19, f24, f02, f01, and f43. These features will be removed and not used in the classification process. The information gain technique produces fewer features than the GR technique. This means that the information gain technique is able to reduce irrelevant features according to the threshold.

Selected features are used as final features as there is no further feature removal, and these features are used for training as well as the desired test data to measure classifier efficiency. We develop the model using WEKA (version 3.9.2) in this research. The WEKA platform simplifies the construction of several data

analysis techniques and offers a JAVA programming language API [37]. It provides tools for categorizing, regressing, grouping, eliminating superfluous traits, creating association rules, and displaying the dataset.

To assess classification performance, we apply *k*-fold cross-validation. The same data is used to separate the dataset into 'k' subsets. Ten folds of data will be used in this study, with each fold being around the same size. They therefore have ten data subsets. The cross-validation test will employ 9-fold for training and 1-fold for testing for each of the 10 data subsets. Three different classifiers, including the SVM, Naive Bayes, and RFC available in WEKA were used to know which classifier outperformed. The confusion matrix for each characteristic in the dataset is used to evaluate how well the models performed. The results of the accuracy performance comparison of three scenarios can be seen in Table 4.

Table 4. A comparison of accuracy for NPD

Methods	All features (%)	Information gain (%)	GR (%)
RFC	99.93	99.96	100
SVM	99.07	97.68	97.79
Naive Bayes	86.52	91.28	91.34

Based on test findings, it is known that the three ML techniques can accurately predict NPD when the feature selection strategy is used. By applying feature selection techniques, the accuracy value of each ML method has increased. This shows that the presence of features affects the classification results. Feature selection selects several features that are able to provide the best results in classification [38]. From the results of the validation test, the GR technique has the highest accuracy value compared to the others.

For all test scenarios, it has previously been demonstrated that the RFC approach is more accurate. Value the accuracy of the RFC method using all features, information gain, and GR with 99.93%, 99.96%, and 100% respectively. We noticed that the RFC with GR has the best of accuracy with 100%. Meanwhile, the Naive Bayes method has the lowest accuracy value compared to the other methods. Because the Naive Bayes method has constraints in class imbalances that affect the classification results. Due to its sensitivity to class distribution, naive Bayes predictions can occasionally be unsuccessful at predicting minority instances.

Furthermore, a comparison of the error rate values of three scenarios is carried out. The error rate is calculated using the confusion matrix. The results of the error rate comparison can be seen in Figure 3.

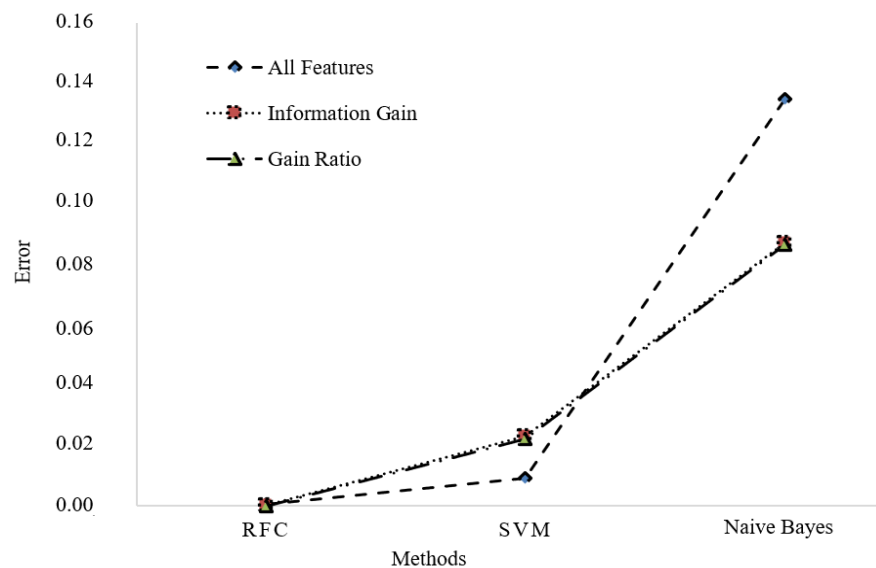


Figure 3. A comparison of the error rate of the methods

Based on Figure 3, the RFC methods have the smallest error rate. The error rate of the RFC using all features, information gain, and GR with 0.07, 0.04, and 0 respectively. In this experiment, we also evaluate the efficiency of the scenarios by measuring the consuming time to develop the model. Consuming time includes the time required to build a predictive model. The comparison of consuming time is shown in Figure 4.

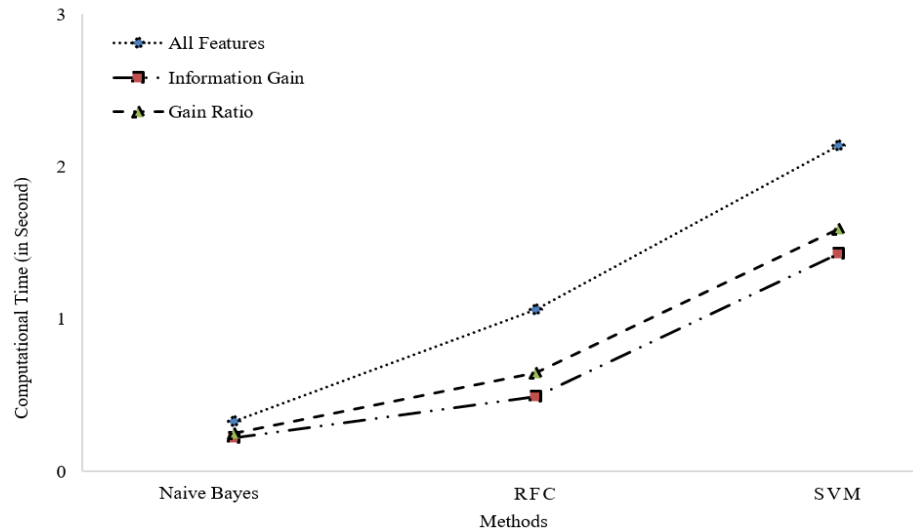


Figure 4. A comparison of consuming time of the methods

Figure 4 presents the consuming time for the ML methods before and after the feature selection using information gain and GR. The test results show that the Naive Bayes with information gain method has the fastest consuming time; it only takes 0.22 seconds. Because Naive Bayes determines the best suitable probability by calculating the likelihood of one class for each existing attribute group. Meanwhile, the SVM approach takes more time than the other two methods. This is because the SVM method's job is more complicated due to using a kernel that seeks out a hyperplane, which makes the processing time longer [39].

In addition, the test results show that the number of features affects the consumption time. The information gain technique produces fewer features than the GR. So, the information gain technique has a faster consuming time when compared to using all the features and GR techniques.

4. CONCLUSION

The test results show that the GR feature selection technique has the highest accuracy value compared to using all features and information gain. In the application of machine learning, the RFC method with a GR has the highest accuracy value of 100%. In the time-consuming test, the naive Bayes method with information gain has the fastest time, which is 0.22 seconds. The test results also show that the number of features analyzed greatly affects processing/consuming time. Therefore, it is necessary to carry out further research to improve feature selection performance to produce relevant and important features.

ACKNOWLEDGMENTS

The authors would like to thank the Universitas Teknokrat Indonesia for providing support in completing this research. This research was funded by LPPM Universitas Teknokrat Indonesia, Research Grants HiPuTS with number 023/UTI/LPPMI/E.1.1/VIII/2023.

REFERENCES




- [1] M. Khazbak, Z. Wael, Z. Ehab, M. Gerorge, and E. Eliwa, "MindTime: Deep Learning Approach for Borderline Personality Disorder Detection," *2021 Int. Mobile, Intelligent, Ubiquitous Comput. Conf. MIUCC 2021*, vol. 21, no. 01, pp. 337–344, 2021, doi: 10.1109/MIUCC52538.2021.9447620.
- [2] T. Jain, A. Jain, P. S. Hada, H. Kumar, V. K. Verma, and A. Patni, "Machine Learning Techniques for Prediction of Mental Health," in *The 3rd International Conference on Inventive Research in Computing Applications (ICIRCA) 2021*, 2021, pp. 1606–1613, doi: 10.1109/ICIRCA51532.2021.9545061.
- [3] C. P. B. Randa and A. E. Permanasari, "Development of Diagnosis Expert System for Personality Disorders," in *Makassar International Conference on Electrical Engineering and Informatics (MICEEI)*, 2014, pp. 180–183, doi: 10.1109/MICEEI.2014.7067335.
- [4] Geneva, "WHO highlights urgent need to transform mental health and mental health care," *WHO int*, 2022. [Online]. Available: [https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care#:~:text=In 2019%2C nearly a billion,6 years lived with disability. \(accessed: 7 May, 2023\).](https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care#:~:text=In 2019%2C nearly a billion,6 years lived with disability. (accessed: 7 May, 2023).)
- [5] N. Hartini, N. A. Fardana, A. D. Ariana, and N. D. Wardana, "Stigma Toward People with Mental Health Problems in Indonesia," *Psychol. Res. Behav. Manag.*, vol. 11, no. 1, pp. 535–541, 2018, doi: 10.2147/PRBM.S175251.
- [6] C. Feng *et al.*, "Individualized Prediction of Trait Narcissism from Whole-Brain Resting-State Functional Connectivity," *Hum.*

- Brain Mapp.*, vol. 39, no. 9, pp. 1–12, 2018, doi: 10.1002/hbm.24205.
- [7] P. Mitra and D. Fluyau, “Narcissistic Personality Disorder,” *National Library of Medicine*, 2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK556001/>.
 - [8] D. S. Stolz, A. Vater, B. H. Schott, S. Roepke, F. M. Paulus, and S. Krach, “Reduced frontal cortical tracking of conflict between self-beneficial versus prosocial motives in Narcissistic Personality Disorder,” *NeuroImage Clin.*, vol. 32, no. 4, pp. 1–11, 2021, doi: 10.1016/j.nicl.2021.102800.
 - [9] L. Faccini and C. S. Allely, “Mass violence in individuals with Autism Spectrum Disorder and Narcissistic Personality Disorder: A case analysis of Anders Breivik using the ‘Path to Intended and Terroristic Violence’ model,” *Aggress. Violent Behav.*, vol. 31, no. 1, pp. 229–236, 2016, doi: 10.1016/j.avb.2016.10.002.
 - [10] M. Silvana, R. Akbar, Derisma, M. Audina, and Firdaus, “Development of Classification Features of Mental Disorder Characteristics Using The Fuzzy Logic Mamdani Method,” *2018 International Conference on Information Technology Systems and Innovation (ICITSI)*, Bandung, Indonesia, 2018, pp. 410–414, doi: 10.1109/ICITSI.2018.8696043.
 - [11] U. S. Reddy, A. V. Thota, and A. Dharun, “Machine Learning Techniques for Stress Prediction in Working Employees,” in *2018 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2018*, 2018, pp. 17–20, doi: 10.1109/ICCIC.2018.8782395.
 - [12] K. Vaishnavi, U. N. Kamath, B. A. Rao, and N. V. S. Reddy, “Predicting Mental Health Illness using Machine Learning Algorithms,” in *Journal of Physics: Conference Series: AICECS*, 2022, pp. 1–7, doi: 10.1088/1742-6596/2161/1/012021.
 - [13] J. F. Dipnall, “Fusing Data Mining, Machine Learning and Traditional Statistics to Detect Biomarkers Associated with Depression,” *Data Mine Mach. Learn Biomarkers Depress.*, vol. 11, no. 2, pp. 1–23, 2016, doi: 10.1371/journal.pone.0148195.
 - [14] M. Soori, B. Arezoo, and R. Dastres, “Sustainable Manufacturing and Service Economics,” *Sustain. Manuf. Serv. Econ.*, vol. 19, no. 1, pp. 1–12, 2023, doi: 10.1016/j.smse.2023.100009.
 - [15] J. Chung and J. Teo, “Mental Health Prediction Using Machine Learning : Taxonomy, Applications, and Challenges,” *Appl. Comput. Intell. Soft Comput.*, vol. 14, no. 1, pp. 1–19, 2022, doi: 10.1155/2022/9970363.
 - [16] J. Wang, S. Zhou, Y. Yi, and J. Kong, “An Improved Feature Selection Based on Effective Range for Classification,” *Sci. World J.*, vol. 15, no. 1, pp. 1–9, 2014, doi: 10.1155/2014/972125.
 - [17] A. Taha, B. Cosgrave, and S. McKeever, “Using Feature Selection with Machine Learning for Generation of Insurance Insights,” *Appl. Sci.*, vol. 12, no. 6, pp. 1–16, 2022, doi: 10.3390/app12063209.
 - [18] J. Nalić, G. Martinović, and D. Žagar, “New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers,” *Adv. Eng. Informatics*, vol. 45, no. 8, pp. 1–9, 2020, doi: 10.1016/j.aei.2020.101130.
 - [19] J. Cai, J. Luo, S. Wang, and S. Yang, “Feature selection in machine learning: A new perspective,” *Neurocomputing*, vol. 300, no. 7, pp. 70–79, 2018, doi: 10.1016/j.neucom.2017.11.077.
 - [20] N. Ansari, “A Survey on Feature Selection Techniques using Evolutionary Algorithms,” *Iraqi J. Sci.*, vol. 62, no. 8, pp. 2796–2812, 2021, doi: 10.24996/ijs.2021.62.8.32.
 - [21] H. Takci, “Improvement of Heart Attack Prediction by the Feature Selection Methods,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 26, no. 1, pp. 1–10, 2018, doi: 10.3906/elk-1611-235.
 - [22] S. Bashir, Z. S. Khan, F. H. Khan, A. Anjum, and K. Bashir, “Improving Heart Disease Prediction Using Feature Selection Approaches,” in *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 2019, pp. 619–623, doi: 10.1109/IBCAST.2019.8667106.
 - [23] K. Verma, B. K. Singh, P. Tripathi, and A. S. Thoke, “Review of Feature Selection Algorithms for Breast Cancer Ultrasound Image,” in *International Conference on New Trends in Intelligent Information and Database Systems*, 2015, pp. 23–32, doi: 10.1007/978-3-319-16211-9.
 - [24] T. A. Alhaj, “Feature Selection Using Information Gain for Improved Structural-Based Alert Correlation,” *PLoS One*, vol. 11, no. 11, pp. 1–11, 2016, doi: 10.1371/journal.pone.0166017.
 - [25] S. Aljawarneh, M. Aldwairi, and M. Bani, “Anomaly-based Intrusion Detection System through Feature Selection Analysis and Building Hybrid Efficient Model,” *J. Comput. Sci.*, vol. 25, no. 1, pp. 152–160, 2018, doi: 10.1016/j.jocs.2017.03.006.
 - [26] R. P. Priyadarsini, M. L. Valarmathi, and S. Sivakumari, “Gain Ratio Based Feature Selection Method for Privacy Preservation,” *ICTACT J. Soft Comput.*, vol. 01, no. 04, pp. 201–205, 2011, doi: 10.21917/ijsc.2011.0031.
 - [27] A. Tamilmani and M. Sughasiny, “Gain Ratio With Optimization Based Feature Selection Method,” *Webology*, vol. 18, no. 6, pp. 6545–6557, 2021.
 - [28] Z. Yang, “Robust Discriminant Feature Selection Via Joint L2,1-Norm Distance Minimization and Maximization,” *Knowledge-Based Syst.*, vol. 207, no. 11, pp. 1–30, 2020, doi: 10.1016/j.knosys.2020.106090.
 - [29] M. I. Prasetyowati, N. U. Maulidevi, and K. Surendro, “Determining Threshold Value on Information Gain Feature Selection to Increase Speed and Prediction Accuracy of Random Forest,” *J. Big Data*, vol. 8, no. 1, pp. 1–22, 2021, doi: 10.1186/s40537-021-00472-4.
 - [30] N. Calik, M. A. Belen, and P. Mahouti, “Deep Learning Base Modified MLP Model for Precise Scattering Parameter Prediction of Capacitive Feed Antenna,” *Spec. Issue Adv. Solut. Methods Model. Complex Electromagn. Probl.*, vol. 1, no. 1, pp. 1–16, 2019, doi: 10.1002/jnm.2682.
 - [31] S. J. Yim, “The utility of smartphone-based, ecological momentary assessment for depressive symptoms,” *J. Affect. Disord.*, vol. 274, no. 9, pp. 602–609, 2020, doi: 10.1016/j.jad.2020.05.116.
 - [32] A. K. Das, S. Goswami, A. Chakrabarti, and B. Chakraborty, “A New Hybrid Feature Selection Approach using Feature Association Map for Supervised and Unsupervised Classification,” *Expert Syst. Appl.*, vol. 88, no. 12, pp. 81–94, 2017, doi: 10.1016/j.eswa.2017.06.032.
 - [33] P. Bhat and K. Dutta, “A Multi-tiered Feature Selection Model for Android Malware Detection Based on Feature Discrimination and Information Gain,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 1–14, 2022, doi: 10.1016/j.jksuci.2021.11.004.
 - [34] M. Trabelsi, N. Meddouri, and M. Maddouri, “A New Feature Selection Method for Nominal Classifier based on Formal Concept Analysis,” in *the 21st International Conference on Knowledge - Based and Intelligent Information & Engineering Systems (KES 2017)*, 2017, pp. 186–194, doi: 10.1016/j.procs.2017.08.227.
 - [35] H. Sulistiani and A. Tjahyanto, “Comparative Analysis of Feature Selection Method to Predict Customer Loyalty,” *IPTEK J. Eng.*, vol. 3, no. 1, pp. 1–5, 2017, doi: 10.12962/joe.v3i1.2257.
 - [36] T. Yan, S. Shen, A. Zhou, and X. Chen, “Prediction of geological characteristics from shield operational parameters by integrating grid search and K-fold cross validation into stacking classification algorithm,” *J. Rock Mech. Geotech. Eng.*, pp. 1–12, 2022, doi: 10.1016/j.jrmge.2022.03.002.
 - [37] K. Kangra and J. Singh, “Comparative analysis of predictive machine learning algorithms for diabetes mellitus,” *Bull. Electr. Eng. Informatics*, vol. 12, no. 3, pp. 1728–1737, 2023, doi: 10.11591/eei.v12i3.4412.




- [38] S. J. Sushma, T. A. Assegie, D. C. Vinutha, and S. Padmashree, "An improved feature selection approach for chronic heart disease detection," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 6, pp. 3501–3506, 2021, doi: 10.11591/eei.v10i6.3001.
- [39] R. Syahputra, G. J. Yanris, and D. Irmayani, "SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter," *Sink. J. dan Penelit. Tek. Inform.*, vol. 7, no. 2, pp. 671–678, 2022, doi: 10.33395/sinkron.v7i2.11430.

BIOGRAPHIES OF AUTHORS






Heni Sulistiani    is a doctoral student at the Faculty of Mathematics and Natural Sciences, Lampung University, Indonesia. She is a lecturer at the Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Lampung, Indonesia. She completed a Bachelor of Informatics Engineering in STMIK Teknokrat in 2012. She holds a Master of Information Systems from Institut Teknologi Sepuluh Nopember (ITS), Surabaya in 2016. She mainly researches data mining, machine learning, and software engineering. She can be contacted at email: henisulistiani@teknokrat.ac.id.






Admi Syarif    is a Professor at the Department of Computer Science, Faculty of Mathematics and Natural Sciences, Lampung University, Indonesia. He received a Bachelor of Science in Mathematics from Padjadjaran University, Indonesia, in 1990 and his Ph.D. in Industrial and Information System Engineering from Ashikaga Institute of Technology, Japan, in 2004. He was a director of the research center of Lampung University from 2010 to 2016; and has been a national research reviewer of the Ministry of Education and Culture, Republic Indonesia, since 2014. He received an award for the best researcher in Information Technology from the Ministry of Research and Technology, Indonesia, in 2009. His research interests include artificial intelligence, mathematical programming, combinatorial optimization, genetic algorithms, and data science. He has presented his papers at several international conferences in Japan (Tokyo, Kyoto, Fukuoka, and so on), the USA (Los Angeles, San Francisco, Las Vegas, Florida, Hawaii), Hong Kong, Australia (Canberra, Sydney, Melbourne, and Lismore), Singapore, Malaysia, Beijing, Zhiang Jia-Jie, Greece, Canada, South Korea (Soul and Jeju), Taiwan, Ireland, and so on. One of his papers appeared in the book fuzzy set and system by Springer Verlag. He also wrote several books in Indonesian. He has been one of the reviewers for the European Journal of Operational Research and Central European Journal of Operational Research and an Asia area editor for JIM. He is a member of the Indonesian Mathematical Society (Indo-MS), Indonesian Operation Research Association (IORA), Association of Computer University (APTİKOM), and Indonesian Lecturer Society (IDS). He can be contacted at email: admi.syarif@fmipa.unila.ac.id.



Kurnia Muludi    is a Doctor of Computer Science. He received a Bachelor of Computer Science from Lampung University, Indonesia, in 1987. He holds a Master in State University of Ghent, Belgium, in 1994 and his Ph.D. in computer science from Institut Teknologi Bandung, Indonesia, in 2011. He mainly researches artificial intelligence, information retrieval, GIS, Information extraction, and bussiness intelligence. He can be contacted at email: kmuludi@fmipa.unila.ac.id.



Warsito    is a Professor at the Department of Physics, Faculty of Mathematics and Natural Science, Lampung University. He completed his bachelor's degree at the Department of Physics, Brawijaya University, Malang, in 1995. Meanwhile, his master's degree was completed in 2000 at Universite De Caen, France. For doctoral education, he completed in 2004 at the same university in France. His other experience is as the Education and Culture Attache at the Indonesian Embassy in Paris. He serves as Deputy for Coordinating Education Quality Improvement and Religious Moderation. He can be contacted at email: warsito@fmipa.unila.ac.id.